

0580
0417

OIIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/995,749A

DATE: 04/22/2002
TIME: 17:19:02

Input Set : A:\Bo438cip.app
Output Set: N:\CRF3\04222002\I995749A.raw

3 <110> APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
4 DIJKHUIZEN, LUBBERT
5 RAHAOUI, HAKIM
6 LEER, ROBERT-JAN

8 <120> TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES

10 <130> FILE REFERENCE: BO43388-CIP

12 <140> CURRENT APPLICATION NUMBER: 09/995,749A

13 <141> CURRENT FILING DATE: 2001-11-29

15 <150> PRIOR APPLICATION NUMBER: 09/604,957

16 <151> PRIOR FILING DATE: 2000-06-28

18 <150> PRIOR APPLICATION NUMBER: EPO 00201871.1

19 <151> PRIOR FILING DATE: 2000-05-25

21 <160> NUMBER OF SEQ ID NOS: 19

23 <170> SOFTWARE: PatentIn Ver. 2.1

25 <210> SEQ ID NO: 1

26 <211> LENGTH: 6026

27 <212> TYPE: DNA

28 <213> ORGANISM: Lactobacillus reuteri

30 <220> FEATURE:

31 <221> NAME/KEY: CDS

32 <222> LOCATION: (161)..(5503)

34 <400> SEQUENCE: 1

35 atacatattt tgggcttctt tttttgttta aaactgtaaa tttgaatttt atttgaaaaa 60
37 attttggcta gaatttgaaa attccctttg aaaaaataaa acatcatagt attataatac 120
39 cgataatcaa attgtttatt ttgatatgaa ggagattaaa atg gaa ata aag aaa 175
40 Met Glu Ile Lys Lys
41 1 5
43 cat ttt aag ttg tac aaa agc ggc aaa caa tgg gta aca gca gca gtg 223
44 His Phe Lys Leu Tyr Lys Ser Gly Lys Gln Trp Val Thr Ala Ala Val
45 10 15 20
47 gct acc gtt gcc gtt tca act gcg ctt ctt tac gga gga gtt gcg cat 271
48 Ala Thr Val Ala Val Ser Thr Ala Leu Leu Tyr Gly Gly Val Ala His
49 25 30 35
51 gct gac caa caa gtt cag caa gct tcc acg act caa gac caa act tct 319
52 Ala Asp Gln Gln Val Gln Gln Ala Ser Thr Thr Gln Asp Gln Thr Ser
53 40 45 50
55 acc gta aat aat gat act gat aaa aca gta gct tta gat act aat act 367
56 Thr Val Asn Asn Asp Thr Asp Lys Thr Val Ala Leu Asp Thr Asn Thr
57 55 60 65
59 gac cag tca gct caa aca act gat aaa aaa caa gta gta tca aat act 415
60 Asp Gln Ser Ala Gln Thr Thr Asp Lys Lys Gln Val Val Ser Asn Thr
61 70 75 80 85
63 aac caa agc aaa act gat gac act tca aca gct gat aag aat tct act 463

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/995,749A

DATE: 04/22/2002

TIME: 17:19:02

Input Set : A:\Bo438cip.app

Output Set: N:\CRF3\04222002\I995749A.raw

64	Asn	Gln	Ser	Lys	Thr	Asp	Asp	Thr	Ser	Thr	Ala	Asp	Lys	Asn	Ser	Thr	
65					90					95				100			
67	tct	aca	cct	gtt	tct	gtt	ttg	cca	tct	aat	aat	act	gaa	aaa	caa	gct	511
68	Ser	Thr	Pro	Val	Ser	Val	Leu	Pro	Ser	Asn	Asn	Thr	Glu	Lys	Gln	Ala	
69					105					110				115			
71	aaa	aat	tat	aat	gag	caa	gac	aaa	gga	aac	tat	ggg	aat	att	gat	act	559
72	Lys	Asn	Tyr	Asn	Glu	Gln	Asp	Lys	Gly	Asn	Tyr	Gly	Asn	Ile	Asp	Thr	
73					120					125				130			
75	gct	tac	ttt	agc	aat	aat	caa	ttg	cat	gtt	tca	gga	tggt	aat	gca	acg	607
76	Ala	Tyr	Phe	Ser	Asn	Asn	Gln	Leu	His	Val	Ser	Gly	Trp	Asn	Ala	Thr	
77					135					140				145			
79	aac	gca	tct	caa	gga	aca	aac	agt	cga	caa	atc	att	gtg	cgt	gat	atc	655
80	Asn	Ala	Ser	Gln	Gly	Thr	Asn	Ser	Arg	Gln	Ile	Ile	Val	Arg	Asp	Ile	
81	150					155				160							
83	aca	acc	aat	aat	gaa	tta	ggt	cgc	act	gat	gta	aca	aac	aat	gtt	gca	703
84	Thr	Thr	Asn	Asn	Glu	Leu	Gly	Arg	Thr	Asp	Val	Thr	Asn	Asn	Val	Ala	
85					170					175					180		
87	cgc	cca	gac	gtt	aag	aat	gtt	cat	aat	gtt	tat	aac	gct	gat	aat	tct	751
88	Arg	Pro	Asp	Val	Lys	Asn	Val	His	Asn	Val	Tyr	Asn	Ala	Asp	Asn	Ser	
89					185					190				195			
91	gga	ttt	gat	gtt	aat	gtc	aat	att	gac	ttt	agc	aag	atg	aaa	gat	tat	799
92	Gly	Phe	Asp	Val	Asn	Val	Asn	Ile	Asp	Phe	Ser	Lys	Met	Lys	Asp	Tyr	
93					200					205				210			
95	cgg	gat	tca	att	gaa	att	gtt	agt	cga	tac	agt	gga	aac	ggt	aaa	tct	847
96	Arg	Asp	Ser	Ile	Glu	Ile	Val	Ser	Arg	Tyr	Ser	Gly	Asn	Gly	Lys	Ser	
97					215					220				225			
99	gtt	gac	tggt	tggt	tcc	caa	ccg	atc	act	ttt	gac	aaa	aac	aac	tat	gct	895
100	Val	Asp	Trp	Trp	Ser	Gln	Pro	Ile	Thr	Phe	Asp	Lys	Asn	Asn	Tyr	Ala	
101	230					235				240				245			
103	tat	ctt	gat	aca	ttt	gaa	gtg	aaa	aat	ggc	gaa	tta	cat	gca	acc	gga	943
104	Tyr	Leu	Asp	Thr	Phe	Glu	Val	Lys	Asn	Gly	Glu	Leu	His	Ala	Thr	Gly	
105					250					255				260			
107	tggt	aat	gct	act	aat	agt	gcg	att	aac	tat	aat	cac	cat	ttt	gtg	att	991
108	Trp	Asn	Ala	Thr	Asn	Ser	Ala	Ile	Asn	Tyr	Asn	His	His	Phe	Val	Ile	
109					265					270				275			
111	ttg	ttt	gat	caa	acg	aat	ggt	aaa	gaa	gta	gca	cga	caa	gaa	gtt	cgt	1039
112	Leu	Phe	Asp	Gln	Thr	Asn	Gly	Lys	Glu	Val	Ala	Arg	Gln	Glu	Val	Arg	
113					280					285				290			
115	gaa	ggt	caa	tca	cgc	cca	gat	gtt	gct	aag	gta	tat	cca	caa	gta	gtt	1087
116	Glu	Gly	Gln	Ser	Arg	Pro	Asp	Val	Ala	Lys	Val	Tyr	Pro	Gln	Val	Val	
117					295					300				305			
119	ggt	gct	gcc	aac	tca	ggc	ttt	aat	gtg	aca	ttt	aat	atc	agt	gat	tta	1135
120	Gly	Ala	Ala	Asn	Ser	Gly	Phe	Asn	Val	Thr	Phe	Asn	Ile	Ser	Asp	Leu	
121	310					315				320				325			
123	gat	tat	act	cac	cag	tac	caa	gtt	ctt	agt	cgt	tac	agc	aat	tct	gat	1183
124	Asp	Tyr	Thr	His	Gln	Tyr	Gln	Val	Leu	Ser	Arg	Tyr	Ser	Asn	Ser	Asp	
125					330					335				340			
127	aat	ggc	gaa	ggt	gat	aac	gtt	acc	tac	tggt	ttt	aat	cca	caa	tcc	att	1231
128	Asn	Gly	Glu	Gly	Asp	Asn	Val	Thr	Tyr	Trp	Phe	Asn	Pro	Gln	Ser	Ile	

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/995,749A

DATE: 04/22/2002
 TIME: 17:19:02

Input Set : A:\Bo438cip.app
 Output Set: N:\CRF3\04222002\I995749A.raw

129		345		350		355		
131	gct cct gct aat caa agt aac cag ggt tat cta gac tca ttt gat att							1279
132	Ala Pro Ala Asn Gln Ser Asn Gln Gly Tyr Leu Asp Ser Phe Asp Ile							
133		360		365		370		
135	agt aaa aat ggt gaa gta aca gta act gga tgg aac gct act gac ttg							1327
136	Ser Lys Asn Gly Glu Val Thr Val Thr Gly Trp Asn Ala Thr Asp Leu							
137		375		380		385		
139	tca gaa tta caa aac aac cat tat gtg att cta ttt gat cag aca gca							1375
140	Ser Glu Leu Gln Asn Asn His Tyr Val Ile Leu Phe Asp Gln Thr Ala							
141	390		395		400		405	
143	ggc aaa caa gtt gca tct gct aaa gct gat tta att tca cgt cca gat							1423
144	Gly Lys Gln Val Ala Ser Ala Lys Ala Asp Leu Ile Ser Arg Pro Asp							
145		410		415		420		
147	gtt gct aaa gct tat cca aca gta aaa aca gct aca aat tct ggc ttc							1471
148	Val Ala Lys Ala Tyr Pro Thr Val Lys Thr Ala Thr Asn Ser Gly Phe							
149		425		430		435		
151	aag gta aca ttt aag gtt aat aac tta caa ccg ggt cac caa tac agc							1519
152	Lys Val Thr Phe Lys Val Asn Asn Leu Gln Pro Gly His Gln Tyr Ser							
153		440		445		450		
155	gtt gta agt cgt ttc tct gcc gat gaa aat ggt aat ggt aat gat aag							1567
156	Val Val Ser Arg Phe Ser Ala Asp Glu Asn Gly Asn Gly Asn Asp Lys							
157		455		460		465		
159	cgc cat aca gat tac tgg ttt agt cca gta ata tta aac cag act gct							1615
160	Arg His Thr Asp Tyr Trp Phe Ser Pro Val Ile Leu Asn Gln Thr Ala							
161	470		475		480		485	
163	tca aac att gat act att aca atg aca tct aat ggt tta cat att gca							1663
164	Ser Asn Ile Asp Thr Ile Thr Met Thr Ser Asn Gly Leu His Ile Ala							
165		490		495		500		
167	ggt tgg atg gca agt gat aac tca att aat gaa aca act cca tac gct							1711
168	Gly Trp Met Ala Ser Asp Asn Ser Ile Asn Glu Thr Thr Pro Tyr Ala							
169		505		510		515		
171	att atc ctc aat aat gga aaa gaa gtt act cgt caa aag atg agc tta							1759
172	Ile Ile Leu Asn Asn Gly Lys Glu Val Thr Arg Gln Lys Met Ser Leu							
173		520		525		530		
175	acc gcc cgt cca gat gta gca gca gta tat cct tca ctt tat aat agt							1807
176	Thr Ala Arg Pro Asp Val Ala Ala Val Tyr Pro Ser Leu Tyr Asn Ser							
177		535		540		545		
179	gct gtt agt ggt ttt gac act act att aaa ttg act aat gat caa tat							1855
180	Ala Val Ser Gly Phe Asp Thr Thr Ile Lys Leu Thr Asn Asp Gln Tyr							
181	550		555		560		565	
183	caa gcg ctt aat ggc caa tta caa gta ttg tta cgt ttt tct aaa gct							1903
184	Gln Ala Leu Asn Gly Gln Leu Gln Val Leu Leu Arg Phe Ser Lys Ala							
185		570		575		580		
187	gct gat ggt aat cca agt ggt gat aat act gta act gat caa ttt agt							1951
188	Ala Asp Gly Asn Pro Ser Gly Asp Asn Thr Val Thr Asp Gln Phe Ser							
189		585		590		595		
191	aaa aat tat gca act act ggt gga aac ttt gac tat gta aaa gta aat							1999
192	Lys Asn Tyr Ala Thr Thr Gly Gly Asn Phe Asp Tyr Val Lys Val Asn							
193		600		605		610		

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/995,749A

DATE: 04/22/2002
 TIME: 17:19:02

Input Set : A:\Bo438cip.app
 Output Set: N:\CRF3\04222002\I995749A.raw

195	ggt aat caa gtt gaa ttt agt ggt tgg cac gca act aac caa tca aat	2047
196	Gly Asn Gln Val Glu Phe Ser Gly Trp His Ala Thr Asn Gln Ser Asn	
197	615 620 625	
199	gat aaa gat tca caa tgg att att gtt tta gtt aat ggt aag gaa gta	2095
200	Asp Lys Asp Ser Gln Trp Ile Ile Val Leu Val Asn Gly Lys Glu Val	
201	630 635 640 645	
203	aag cgt caa tta gtt aat gat act aaa gag gga gct gct ggc ttc aac	2143
204	Lys Arg Gln Leu Val Asn Asp Thr Lys Glu Gly Ala Ala Gly Phe Asn	
205	650 655 660	
207	cga aac gat gtc tac aaa gta aat cca gct att gaa aac agt tct atg	2191
208	Arg Asn Asp Val Tyr Lys Val Asn Pro Ala Ile Glu Asn Ser Ser Met	
209	665 670 675	
211	tct gga ttc caa ggc att att act tta cct gtg aca gtt aaa aac gaa	2239
212	Ser Gly Phe Gln Gly Ile Ile Thr Leu Pro Val Thr Val Lys Asn Glu	
213	680 685 690	
215	aat gtc caa ctt gtt cat cgg ttt agt aac gat gtg aag act ggt gaa	2287
216	Asn Val Gln Leu Val His Arg Phe Ser Asn Asp Val Lys Thr Gly Glu	
217	695 700 705	
219	ggt aac tat gtt gat ttc tgg tca gag cta atg cct gtt aag gat agc	2335
220	Gly Asn Tyr Val Asp Phe Trp Ser Glu Leu Met Pro Val Lys Asp Ser	
221	710 715 720 725	
223	ttc caa aag ggg aat ggc cca ctt aag caa ttt ggc tta caa act att	2383
224	Phe Gln Lys Gly Asn Gly Pro Leu Lys Gln Phe Gly Leu Gln Thr Ile	
225	730 735 740	
227	aac ggt caa caa tat tat att gac cca aca act ggt caa cca cgt aag	2431
228	Asn Gly Gln Gln Tyr Tyr Ile Asp Pro Thr Thr Gly Gln Pro Arg Lys	
229	745 750 755	
231	aat ttc tta tta caa agt gga aat aat tgg att tac ttt gat agt gat	2479
232	Asn Phe Leu Leu Gln Ser Gly Asn Asn Trp Ile Tyr Phe Asp Ser Asp	
233	760 765 770	
235	act ggt gtg ggt act aat gca ctt gaa tta caa ttt gca aag gga act	2527
236	Thr Gly Val Gly Thr Asn Ala Leu Glu Leu Gln Phe Ala Lys Gly Thr	
237	775 780 785	
239	gtt tca tct aat gaa caa tac cgt aac ggt aat gca gct tac agt tat	2575
240	Val Ser Ser Asn Glu Gln Tyr Arg Asn Gly Asn Ala Ala Tyr Ser Tyr	
241	790 795 800 805	
243	gat gac aag agt atc gaa aat gta aat ggt tac tta aca gca gat aca	2623
244	Asp Asp Lys Ser Ile Glu Asn Val Asn Gly Tyr Leu Thr Ala Asp Thr	
245	810 815 820	
247	tgg tac cgt cca aaa cag atc tta aag gat gga act acc tgg act gac	2671
248	Trp Tyr Arg Pro Lys Gln Ile Leu Lys Asp Gly Thr Thr Trp Thr Asp	
249	825 830 835	
251	tca aaa gaa aca gat atg cga cca atc ttg atg gta tgg tgg cct aat	2719
252	Ser Lys Glu Thr Asp Met Arg Pro Ile Leu Met Val Trp Trp Pro Asn	
253	840 845 850	
255	act ctt acc caa gca tac tac ctt aat tac atg aaa caa cat ggt aat	2767
256	Thr Leu Thr Gln Ala Tyr Tyr Leu Asn Tyr Met Lys Gln His Gly Asn	
257	855 860 865	
259	tta tta cca tct gct tta cca ttc ttt aat gcg gat gct gat cct gca	2815

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/995,749A

DATE: 04/22/2002

TIME: 17:19:02

Input Set : A:\Bo438cip.app

Output Set: N:\CRF3\04222002\I995749A.raw

```

260 Leu Leu Pro Ser Ala Leu Pro Phe Phe Asn Ala Asp Ala Asp Pro Ala
261 870 875 880 885
263 gaa tta aat cat tat tcc gaa att gtg caa caa aat att gaa aaa cga 2863
264 Glu Leu Asn His Tyr Ser Glu Ile Val Gln Gln Asn Ile Glu Lys Arg
265 890 895 900
267 att agt gaa acc gga aat act gat tgg tta cgt act tta atg cac gat 2911
268 Ile Ser Glu Thr Gly Asn Thr Asp Trp Leu Arg Thr Leu Met His Asp
269 905 910 915
271 ttt gtt act aac aat ccg atg tgg aat aag gat agt gaa aat gtt aac 2959
272 Phe Val Thr Asn Asn Pro Met Trp Asn Lys Asp Ser Glu Asn Val Asn
273 920 925 930
275 ttt agt ggt att caa ttc caa ggc gga ttc tta aag tat gaa aac tca 3007
276 Phe Ser Gly Ile Gln Phe Gln Gly Gly Phe Leu Lys Tyr Glu Asn Ser
277 935 940 945
279 gat tta acg cct tat gct aac tct gat tat cgc tta ctt ggt cgg atg 3055
280 Asp Leu Thr Pro Tyr Ala Asn Ser Asp Tyr Arg Leu Leu Gly Arg Met
281 950 955 960 965
283 cca atc aat att aag gat caa aca tat cgg gga caa gaa ttc cta ctt 3103
284 Pro Ile Asn Ile Lys Asp Gln Thr Tyr Arg Gly Gln Glu Phe Leu Leu
285 970 975 980
287 gct aac gat att gat aac tct aat cct gtt gtt caa gca gaa caa tta 3151
288 Ala Asn Asp Ile Asp Asn Ser Asn Pro Val Val Gln Ala Glu Gln Leu
289 985 990 995
291 aac tgg tta tac tat ctc ttg aac ttt gga acg atc aca gct aat aat 3199
292 Asn Trp Leu Tyr Tyr Leu Leu Asn Phe Gly Thr Ile Thr Ala Asn Asn
293 1000 1005 1010
295 gat caa gct aat ttt gat tct gta cgg gta gat gca ccg gat aat att 3247
296 Asp Gln Ala Asn Phe Asp Ser Val Arg Val Asp Ala Pro Asp Asn Ile
297 1015 1020 1025
299 gat gcc gat ctt atg aat atc gct cag gac tac ttt aat gct gca tat 3295
300 Asp Ala Asp Leu Met Asn Ile Ala Gln Asp Tyr Phe Asn Ala Ala Tyr
301 1030 1035 1040 1045
303 ggt atg gac tca gat gct gtc tca aat aag cat att aat att ctt gaa 3343
304 Gly Met Asp Ser Asp Ala Val Ser Asn Lys His Ile Asn Ile Leu Glu
305 1050 1055 1060
307 gac tgg aat cat gct gat ccg gaa tac ttt aat aag atc gga aat cca 3391
308 Asp Trp Asn His Ala Asp Pro Glu Tyr Phe Asn Lys Ile Gly Asn Pro
309 1065 1070 1075
311 caa ttg aca atg gat gat act att aag aat tcc ctg aat cat ggg ctt 3439
312 Gln Leu Thr Met Asp Asp Thr Ile Lys Asn Ser Leu Asn His Gly Leu
313 1080 1085 1090
315 tca gat gca act aat cgt tgg gga tta gat gca att gtt cat cag tca 3487
316 Ser Asp Ala Thr Asn Arg Trp Gly Leu Asp Ala Ile Val His Gln Ser
317 1095 1100 1105
319 tta gct gat cgt gaa aat aat tcc acg gaa aat gtt gta att cct aat 3535
320 Leu Ala Asp Arg Glu Asn Asn Ser Thr Glu Asn Val Val Ile Pro Asn
321 1110 1115 1120 1125
323 tac agt ttc gtt cgg gct cac gat aat aat tct caa gat caa att caa 3583
324 Tyr Ser Phe Val Arg Ala His Asp Asn Asn Ser Gln Asp Gln Ile Gln

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/995,749A

DATE: 04/22/2002
TIME: 17:19:03

Input Set : A:\Bo438cip.app
Output Set: N:\CRF3\04222002\I995749A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; N Pos. 9,15,18,21

Seq#:7; N Pos. 2,5,8,17,21

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:17; Line(s) 1454